

<https://helda.helsinki.fi>

Identification of Economic Shocks by Inequality Constraints in Bayesian Structural Vector Autoregression

Lanne, Markku

2020-04

Lanne , M & Luoto , J 2020 , ' Identification of Economic Shocks by Inequality Constraints in Bayesian Structural Vector Autoregression ' , Oxford Bulletin of Economics and Statistics , vol. 82 , no. 2 , pp. 425-452 . <https://doi.org/10.1111/obes.12338>

<http://hdl.handle.net/10138/317213>

<https://doi.org/10.1111/obes.12338>

cc_by

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.



Identification of Economic Shocks by Inequality Constraints in Bayesian Structural Vector Autoregression*

MARKKU LANNE and JANI LUOTO

Faculty of Social Sciences, University of Helsinki, Helsinki, Finland

Abstract

Theories often make predictions about the signs of the effects of economic shocks on observable variables, thus implying inequality constraints on the parameters of a structural vector autoregression (SVAR). We introduce a new Bayesian procedure to evaluate the probabilities of such constraints, and, hence, to validate the theoretically implied economic shocks. We first estimate a SVAR, where the shocks are identified by statistical properties of the data, and subsequently label these statistically identified shocks by the Bayes factors calculated from their probabilities of satisfying given inequality constraints. In contrast to the related sign restriction approach that also makes use of theoretically implied inequality constraints, no restrictions are imposed. Hence, it is possible that only a subset or none of the theoretically implied shocks can be labelled. In the latter case, we conclude that the data do not lend support to the theory implying the signs of the effects in question. We illustrate the method by empirical applications to the crude oil market, and U.S. monetary policy.

I. Introduction

The structural vector autoregressive (SVAR) model is one of the prominent tools in empirical macroeconomics. While the reduced-form VAR is useful for describing the joint dynamics of a number of time series, it is only when some structure is imposed upon it that interesting economic questions apart from forecasting can be addressed. Typically SVAR analysis involves tracing out the dynamic effects (impulse responses) of economic shocks on the variables included in the model, and these shocks are often identified by restricting their effects in various ways (for a comprehensive survey on SVAR models, see Kilian and Lütkepohl, 2017). Recently, identification by sign restrictions, put forth by Faust (1998); Canova and De Nicoló (2002); Uhlig (2005), has become increasingly

JEL Classification numbers: C32, C5, E52, Q41.

*We thank Heino Bohn Nielsen (editor) and three anonymous referees for useful comments. Financial support from the Academy of Finland (grant 308628) is gratefully acknowledged. The first author also acknowledges financial support from CREATES (DNRF78) funded by the Danish National Research Foundation, while the second author is grateful for financial support from the Research Funds of the University of Helsinki.

popular in the macroeconomic SVAR literature. Compared to most other approaches, sign restrictions only constraining the signs of the effects of (some of) the shocks to accord with economic theory or institutional knowledge, are less stringent, yet manage to convey economic intuition. Therefore, they have a great appeal in empirical research.

In this paper, we propose a formal procedure to identify economic shocks without actually imposing any restrictions on the parameters of the SVAR model, while still making use of the signs of the effects of shocks. These signs can be easily expressed in the form of inequality constraints on the parameters of the SVAR model.¹ Our starting point is the SVAR model where, following Hyvärinen *et al.* (2010); Lanne *et al.* (2017), identification is achieved by means of statistical properties of the data. The statistically identified structural shocks (errors) have no economic meaning as such, but for interpretation, they need to be labelled using external information. To that end, sign constraints have been used in the previous statistical identification literature (see, e.g. Herwartz and Lütkepohl, 2014; Lütkepohl and Netsunajev, 2014; Lanne *et al.*, 2017). The idea of this approach is to visually check whether the impulse responses implied by the uniquely identified SVAR model satisfy the constraints. If they are satisfied, the shocks can be labelled accordingly. Our procedure formalizes this approach by quantifying the likelihood of the inequality constraints. It also has the advantage that it uses all information from the joint (posterior) distribution of the estimator of the impulse responses, while the previous approach is based on their (pseudo) marginal sampling distributions. The latter approach is somewhat deficient and unreliable, akin to a joint hypothesis testing using the usual *t* statistics for testing the restrictions one at a time.

Our analysis is based on Bayesian inference that facilitates straightforward assessment of inequality constraints by posterior odds or Bayes factors (see, e.g. Koop, 2003, 39–40). In particular, as shown in section III, each set of inequality constraints implies a different model, whose posterior probability can be interpreted as the probability of the constraints. This probability can then be transformed to the Bayes factor to weigh the posterior evidence against the case where no constraints are imposed (see, e.g. Kass and Raftery, 1995). Hence our approach facilitates the identification of the shocks that are the likeliest to satisfy the constraints (i.e. are the likeliest to be the structural shocks of interest). It may also turn out that only a subset or none of the inequality constraints are supported by the data. It is then concluded that the constraints that the data do not lend support to, are not useful in identifying the economic shocks in question. In this case, an alternative set of constraints, potentially implied by a competing economic theory, could be entertained, or the subsequent analysis may concentrate only on the subset of the shocks that are identified.

A major difference between our approach and imposing sign restrictions is that the latter only achieves set identification (see, e.g. Baumeister and Hamilton, 2015). Therefore, assessing the plausibility of the given inequality constraints (sign restrictions) is not straightforward, whereas, due to point identification, our approach facilitates direct calculation of the Bayes factor for the constrained model against the unconstrained one.

¹ Inequality constraints may also be imposed on other functions of structural parameters, not just on impulse responses, such as historical decompositions considered in Antolín-Díaz and Rubio-Ramírez (2018) (we thank an anonymous referee for pointing this out). Our procedure obviously generalizes in a straightforward manner to such cases.

Nevertheless, a few suggestions concerning the validation of sign restrictions have been put forth in the previous literature. Straub and Peersman (2006) used the proportion of discarded models as an indicator of how well the New Keynesian model that had generated the restrictions, fits the data. This indicator is, however, ambiguous because a high rejection rate may just as well indicate sharp identification (the set of acceptable models is small) or an inefficient sampler as lack of fit. Piffer (2016) formalized this approach, but his procedure seems difficult to generalize beyond the bivariate VAR model. Baumeister and Hamilton (2015) illustrated how the effect of the tightness of priors on the posteriors yields information on the plausibility of the sign restrictions. However, because only set identification is achieved, the posterior will still be driven by the priors. Furthermore, this approach is applicable only when the priors are explicitly spelled out, while, because of point identification, our approach does not even require the use of informative priors. Finally, Giacomini and Kitagawa (2014) suggested reporting the posterior probability of a non-empty identified set as a measure of posterior belief for the plausibility of the imposed sign restrictions. This probability is, of course, one, as long as the given sign restrictions are not impossible in the posterior sense.

We illustrate our method by means of two empirical applications. One of them focuses on the identification of a monetary policy shock in Uhlig's (2005) SVAR model, while the other involves multiple shocks identified by inequality constraints in Kilian's (2009) model of the crude oil market. In Uhlig's model, we find two shocks that satisfy the inequality constraints involved in his sign restrictions. This possibility was also discussed by Uhlig (2005), who was worried that a money demand shock might satisfy the same sign restrictions as the monetary policy shock. If this is the case, the conventional approach to sign restrictions yields a linear combination of the two shocks, while our approach, by construction, produces two separate shocks. Inspection of our impulse responses indeed lends support to this insight. As for Kilian's (2009) model, our procedure managed to successfully identify all three structural shocks with relatively high probability.

The rest of the paper is organized as follows. In section II, we describe the SVAR model and discuss its identification along the lines of Hyvärinen *et al.* (2010); Lanne *et al.* (2017). Section III introduces the procedure for computing the probabilities of the inequality constraints, and the corresponding Bayes factors, and finding the plausible economic shocks among all the statistically identified shocks. In sections Single structural shock and Multiple structural shocks, we propose stepwise procedures for the cases of a single and multiple identified economic shocks, respectively. Section IV contains the two empirical applications discussed above. Section V concludes. Description of the Metropolis-within-Gibbs algorithm used for the estimation of the posterior distribution of the parameters of our fully identified SVAR model is deferred to Appendix A, while in Appendix B, we discuss the computation of impulse responses and forecast error variance decompositions of the identified shocks.

II. Model

Consider the n -variate structural VAR(p) model

$$y_t = a + A_1 y_{t-1} + \cdots + A_p y_{t-p} + B \varepsilon_t, \quad (1)$$

where y_t is a vector of time series of interest, a is an intercept term, A_1, \dots, A_p are $n \times n$ coefficient matrices, and the matrix B summarizing the contemporaneous structural relations of the errors is assumed non-singular. In order to facilitate identification of matrix B , we assume that the error process $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{nt})'$ consists of independent non-Gaussian components. Specifically, following Lanne *et al.* (2017), we make the following assumption.

Assumption 1.

- (i) The error process $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{nt})'$ is a sequence of independent and identically distributed random vectors with each component ε_{it} , $i = 1, \dots, n$, having zero mean and finite positive variance σ_i^2 .
- (ii) The components of $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{nt})$ are (mutually) independent and at most one of them has a Gaussian marginal distribution.²

In the empirical applications in section IV, we assume that each component of the error vector individually follows a t distribution. Because a t -distributed random variable converges to a Gaussian as the number of degrees of freedom approaches infinity, this is more general than the usual (implicit) normality assumption and affords more flexibility (see, for example, Koop, 2003, p. 126) for a more detailed discussion).

If the process y_t is stable, i.e.

$$\det(I_n - A_1 z - \dots - A_p z^p) \neq 0, |z| \leq 1 (z \in \mathbb{C}), \quad (2)$$

the SVAR(p) model (1) has a moving average representation

$$y_t = \mu + \sum_{j=0}^{\infty} \Psi_j B \varepsilon_{t-j}, \quad (3)$$

where μ is the unconditional expectation of y_t , Ψ_0 is the identity matrix, and Ψ_j , $j = 1, 2, \dots$, are obtained recursively as $\Psi_j = \sum_{l=1}^j \Psi_{j-l} A_l$. The k th column of $\Psi_j B \equiv \Theta_j$, $j = 0, 1, \dots$, contains the impulse responses of the k th structural shock ε_{kt} , $k = 1, \dots, n$, and it is these impulse responses that are the main object of interest in SVAR analysis. An integrated VAR(p) process does not satisfy the stability condition above, and hence, has no moving average representation. Nevertheless, the impulse responses are also in this case given by the same recursion (see Lütkepohl, 2005, section 6.7). In the absence of cointegration, the model can alternatively be specified for the difference of y_t . If the elements of y_t are cointegrated, the relevant moving average representation is the multivariate version of the Beveridge-Nelson decomposition of y_t (see Lütkepohl, 2005, proposition 6.1) instead of (3). For simplicity, we assume below that the stability condition (2) holds.

Under the non-Gaussianity and independence assumptions on the error term ε_t above, matrix B is unique apart from permutation and scaling of its columns as shown in the following proposition adapted from Proposition 1 in Lanne *et al.* (2017); for a proof, see Lanne *et al.* (2017, Appendix A).

² It was recently shown by Lanne and Luoto (2019) that the components of the error term ε_t need not be mutually independent for identification, but they may have, for example, mutually dependent conditional variances. However, because our goal is to introduce a procedure for labelling the statistically identified shocks, not to propose a robust estimation method with respect to misspecification of the error distribution, we follow Lanne *et al.* (2017) in assuming that the components of ε_t are mutually independent. In particular, our estimation algorithm described in Appendix A is based on this assumption.

Proposition 1. Assume that, in addition to (3), y_t generated by the SVAR model (1) under the stationarity condition (2) and Assumption 1 has another moving average representation, $y_t = \mu^* + \sum_{j=0}^{\infty} \Psi_j^* B^* \varepsilon_{t-j}^*$, where B^* is a non-singular $n \times n$ matrix, μ^* is an $n \times 1$ vector, $\Psi_0^* = I_n$ and Ψ_j^* ($j = 1, \dots, p$) are $n \times n$ coefficient matrices determined by the recursion $\Psi_j^* = \sum_{l=1}^j \Psi_{j-l}^* A_l^*$ with $A^*(z) = I_n - A^* z_1 - \dots - A_p^* z^p$ satisfying condition (2), and $\varepsilon_t^* = (\varepsilon_{1t}^*, \dots, \varepsilon_{nt}^*)'$ is an error process satisfying Assumption 1. Then for some diagonal matrix D with non-zero diagonal elements, some permutation matrix P , and for all t , $B^* = BDP$, $\varepsilon_t^* = P'^{-1} \varepsilon_t$, $\mu^* = \mu$, and $\Psi_j^* = \Psi_j$ ($j = 0, 1, \dots$).

Proposition 1 characterizes a class of observationally equivalent SVAR processes that differ only with respect to the ordering and scaling of the structural shocks in the vector ε_t . The (rescaled) error vector of any of these $n!$ SVAR processes thus consists of exactly the same elements, whose ordering varies, and each of the SVAR processes produces the same impulse responses. In order to pick one of these observationally equivalent SVAR processes, we employ the identification scheme (described in detail in Step 2 of Algorithm 1 in section Single structural shock) of Lanne *et al.* (2017, section 3.3).³ In this sense, Assumption 2 coupled with the identification scheme guarantees point identification. However, even though we are able to uniquely choose a particular permutation from the set of all the $n!$ permutations, the impulse responses cannot be related to given equations of the process, or interpreted as shocks to given variables in y_t . In other words, although the structural shocks and their impulse responses may be uniquely identified by picking a particular permutation in a unique manner, the shocks cannot be labelled or be given any economic interpretation without additional information that may come in various forms, such as short-run or long-run restrictions, or inequality constraints on the effects of the shocks.

It is instructive to contrast Proposition 1 with the case of identification by sign restrictions assuming Gaussian errors. When the covariance matrix of ε_t is assumed diagonal, we can always replace $B\varepsilon_t$ in (1) by $B^* \varepsilon_t^*$, where $B^* = B\Lambda Q$ and $\varepsilon_t^* = Q^{-1} \Lambda^{-1} \varepsilon_t$ with Q orthogonal and Λ diagonal. In identification by sign restrictions, $B\Lambda$ is typically taken as the lower-triangular Cholesky factor of the covariance matrix of the residuals of the corresponding reduced-form VAR model, and the set of arbitrary orthogonal matrices Q producing impulse responses satisfying the sign restrictions defines the set of identified SVAR models that has a continuum of elements (see, e.g., Rubio-Ramírez, Waggoner and Zha, 2010). In contrast, in the non-Gaussian case, the set of admissible orthogonal matrices Q has only $n!$ elements, each corresponding to one permutation of the structural errors.

Despite the set Q having only $n!$ elements under Assumption 1, point identification is actually achieved. This can be seen by noticing that any permutation of the columns of B produces the same shocks (with a given size determined by the diagonal elements of Λ) and impulse responses. Thus, from the viewpoint of impulse response analysis, it is irrelevant which permutation is chosen, and therefore, the permutation can be fixed. The identification scheme in Step 2 of Algorithm 1 in section Single structural shock, provides a recipe for picking a particular permutation (and scaling matrix) from the set of all the $n!$ permutations. The sequence of transformations described in Step 2, actually constrains the permissible

³ For a proof that this identification scheme indeed yields a unique SVAR process, see Appendix A in Lanne *et al.* (2017). This is not the only way of pinpointing a unique model, but a number of alternative identification schemes have been put forth in the related literature (see Lanne *et al.*, 2017, and the references therein.)

values of the matrix B to the set \mathcal{B} , defined such that if $B, B^* \in \mathcal{B}$, then necessarily $B = B^*$. The fact that this scheme provides a formal solution to the identification problem, is shown in Appendix A of Lanne *et al.* (2017). This insight is central to our entire analysis.

To take a simple example, let us consider the bivariate SVAR (1) model

$$y_t = A_1 y_{t-1} + B \varepsilon_t, \quad (4)$$

where $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ satisfies Assumption 2, and, for simplicity, normalize its covariance matrix to the identity matrix. Furthermore, let us concentrate on positive shocks. Then, (4) is observationally equivalent only to the model where B is replaced by B^* obtained by reversing the order of its columns. In other words, the impact effects of the structural shocks ε_{1t} and ε_{2t} are given, respectively, by the first and second columns of either

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \text{ or } B^* = \begin{pmatrix} b_{12} & b_{11} \\ b_{22} & b_{21} \end{pmatrix}.$$

Hence, according to the definition in Rubio-Ramírez *et al.* (2010), the structural shocks ε_{1t} and ε_{2t} are set identified, with the identified set consisting of two elements. However, it is obvious that the shock ε_{1t} implied by the permutation of the columns in B is identical to ε_{2t} implied by the permutation in B^* , and similarly for the shock ε_{2t} implied by B and ε_{1t} implied by B^* . Because both permutations produce the same shocks (that without further information cannot be given any economic interpretation), it is irrelevant which permutation we choose. As already discussed, Step 2 of Algorithm 1, provides an identification scheme for picking a particular permutation (either B or B^*) from the set of all the $n!$ ($= 2$ here) permutations. Given this scheme, the model is thus point identified.

While the two structural shocks and the parameters b_{11} , b_{12} , b_{21} and b_{22} can be uniquely identified, labelling the shocks calls for additional information, such as inequality constraints from economic theory. Let us, for example, assume that $y_t = (p_t, q_t)'$, with p_t and q_t the price level and quantity, respectively. The bivariate SVAR(1) model then becomes the simple market (demand/supply) model of Fry and Pagan (2011) if one of the shocks (a positive demand shock) has a positive effect on impact on both p_t and q_t , and the other shock (a positive cost (supply) shock) has a positive impact effect on p_t and a negative impact effect on q_t . If both elements in only one column of the estimated impact matrix are positive, the corresponding shock can be labeled the demand shock. Depending on the permutation, the demand shock is either ε_{1t} or ε_{2t} . If the first element in the remaining column is positive and the second element is negative, the other shock can be labelled the cost shock.⁴ It is, of course, possible that these inequality constraints are able to label only one or neither of the point-identified structural shocks in a given data set.

III. Inference on inequality constraints

The discussion in section II made it clear that because under the non-Gaussianity and independence assumptions the impact matrix B in (1) is uniquely identified (apart from permutation and scaling of its columns), also the structural shocks and their impulse

⁴ We have expressed the sign pattern for positive shocks, but it should be borne in mind that any shock with impact effects of the same (opposite) sign on p_t and q_t is a demand (cost) shock.

responses are identified. Furthermore, the only difference between the models corresponding to the different permutations of the columns of B is the ordering of the shocks, and therefore, a model with any fixed permutation facilitates labelling the shocks using inequality constraints obtained, say, from economic theory, as demonstrated by the example in section II.

Inference concerning the structural shocks is based on the posterior distribution of the parameters. Specifically, we first fix the permutation, and then compute the posterior probability of each combination of the shocks being the one that satisfies the inequality constraints. In practice, this is carried out by drawing from the posterior distribution of the relevant parameters and counting the share of draws that do not violate the constraints. Each combination of the shocks satisfying the inequality constraints corresponds to a different (constrained) model, and to weigh the posterior evidence in favour of each constrained model against the unconstrained model, we recommend calculating the Bayes factor from these posterior probabilities. Following Kass and Raftery (1995), in this paper, we deem the evidence substantial if the value of the Bayes factor exceeds 3.2.

To make things concrete, let us consider again the market model example in section II. Suppose, we want to find out which (if either) of the point-identified shocks is the demand shock. As suggested above, we compute the posterior probability of each of the shocks satisfying the inequality constraints related to the demand shock (and the other shock not satisfying them), and then calculate the associated Bayes factors against the unconstrained case. Provided the evidence against the latter is substantial, we label the shock associated with the greater Bayes factor as the demand shock. In other words, we first compute the posterior probability that $b_{11} > 0$ and $b_{21} > 0$ (and b_{12} and b_{22} are not both positive), and then the posterior probability that $b_{12} > 0$ and $b_{22} > 0$ (and b_{11} and b_{21} are not both positive). Since the choice of permutation is irrelevant, let us fix it to B . Then, if the Bayes factor based the former probability is greater than that based on the latter, and greater than 3.2, say, we label ε_{1t} as the demand shock. In the opposite case, ε_{2t} is labelled as the demand shock. If both Bayes factors are smaller than 3.2, we conclude that the inequality constraints considered are not useful in labelling the demand shock.

If both Bayes factors are greater than 3.2 (i.e. both shocks satisfy the same constraints with relatively high posterior probability), inference is more complicated. In that case, we proceed by calculating the Bayes factor comparing the first to the second constrained model. Values greater (smaller) than 3.2 of this Bayes factor facilitate labelling ε_{1t} (ε_{2t}) as the demand shock. However, if also both of these Bayes factors are smaller than 3.2, additional information is needed to discriminate between the plausible shocks.

Our procedure generalizes in a straightforward manner to the case where we want to label both shocks in this example. Then, we first compare the Bayes factor based on the posterior probability of the restriction $b_{11} > 0$, $b_{21} > 0$, $b_{12} > 0$ and $b_{22} < 0$ to the one based on the posterior probability of the restriction $b_{11} > 0$, $b_{21} < 0$, $b_{12} > 0$ and $b_{22} > 0$. If the data lend greater support to the former restriction than the latter (and the associated Bayes factor is greater than 3.2), we label ε_{1t} the demand shock and ε_{2t} the cost shock. In the opposite case, the labels of the shocks are reversed, and if both Bayes factors are smaller than 3.2, we conclude that the inequality constraints are not useful in identifying the shocks in question.

The convention in the sign restriction literature is to express the sign pattern for positive shocks. However, a negative shock, having effects of the opposite sign on all the constrained

variables, of course, also satisfies the inequality constraints (see, e.g., Fry and Pagan, 2011 for a discussion). In other words, the sign patterns determine whether the constraints hold, not the signs as such. To this end, we normalize one of the rows of B such that it corresponds to one of the inequality constraints in all draws from the posterior distribution of B . Because the responses to a negative and a positive shock are symmetrical, we can then base the analysis on either of them. For instance, if in the market model example discussed above, the top row (corresponding to the effects of shocks on p_t) is normalized positive, all shocks corresponding to a column of B with a positive (negative) element in the bottom row would be labelled a demand (cost) shock, provided the other shock does not satisfy the same constraints.

It is important to notice that also in the case of partial identification, we are able to uniquely capture the shocks satisfying the inequality constraints, provided they exist. This is a great benefit over the sign restriction approach, and follows from the fact that under our assumptions the model is point-identified. As also pointed out by Uhlig (2005), without additional restrictions, the shocks captured by the conventional approach may actually be a combination of several economic shocks.

In this section, we concentrate on assessing the inequality constraints and giving the identified shocks economically meaningful labels, while we defer the discussion on the computation of the impulse responses and forecast error variance decompositions to Appendix B. We set out with the case of a single structural shock identified by inequality constraints on only the impact effects, and then proceed to the more general case of constraints on the first $q + 1$ impulse responses. Discussion on the case of multiple structural shocks concludes the section.

Single structural shock

Suppose we are interested in finding the impulse responses of a single shock, the expected signs of whose impact effects on J of the variables in y_t are given. This might be, say, the monetary policy shock with a non-positive impact effect on prices and non-borrowed reserves and a non-negative impact effect on the Federal funds rate (cf. the empirical application in section Single monetary policy shock). Let us collect these inequality constraints in a $J \times n$ matrix R , whose elements equal 1, -1 , or 0, and define a set Q such that

$$Q = \{\theta_{0k} : R\theta_{0k} \geq 0_{J \times 1}\}, \quad (5)$$

where θ_{0k} is the k th column of Θ_0 , or equivalently, of the impact matrix B , corresponding to shock ε_{kt} . The set Q thus consists of the columns of B that satisfy the inequality constraints. Although we are after a single shock, Q may contain multiple columns of B , or it may be empty, i.e. there may be more than one shock or no shock satisfying the constraints. This is in contrast to the conventional approach in the sign restriction literature, where a single shock satisfying the restrictions, by construction, is found.

Since our assumptions only identify B up to permutation of its columns, any (or none) of the n components of ε_t can be the structural shock satisfying the inequality constraints. In order to assess the plausibility of one of the shocks being the shock of interest, we fix the permutation (i.e. choose one of the permutations), and then compute for each shock ε_{kt} , $k = 1, \dots, n$, the conditional probability of being this shock (conditional on the vector of data, y , obtained by stacking y_t for $t = 1, \dots, T$),

$$\Pr(\theta_{0k} \in Q, \theta_{0,m \neq k} \in Q^c | \mathbf{y}), \quad (6)$$

where Q^c denotes the complement of Q , and $m \in \{1, \dots, n\}$. For each $k \in \{1, \dots, n\}$, the quantity (6) can be interpreted as the posterior probability of the constrained SVAR model, where the inequality constraints embodied in R are imposed on the k th column of B only (cf. Koop, 2003, p. 81 in the context of the linear regression model). Among the n models, we expect the one that satisfies the constraints in the (true) data-generating process (DGP) (i.e. the model for which $\theta_{0k} \in Q$ in the DGP) to have a high posterior probability (relatively), while the probabilities of the other models should be close to zero, provided they do not satisfy the same restrictions. As already discussed, (6) can be easily computed by drawing from the posterior distribution of the relevant parameters and counting the share of draws satisfying the constraints defined in (5).

As already discussed, to facilitate interpretation of the posterior model probabilities (6), we recommend computing the associated Bayes factors of each constrained SVAR model against the unconstrained model. The Bayes factor of the constrained model (involving the inequality constraints embodied in R imposed on the k th column of B only) against the unconstrained model is the ratio of the corresponding marginal data densities:

$$\frac{\int_S p(\mathbf{y} | \phi) p(\phi) d\phi}{\int_{\Phi} p(\mathbf{y} | \phi) p(\phi) d\phi} = \frac{\int_S p(\mathbf{y} | \phi) p(\phi) d\phi}{\int_{\Phi} p(\mathbf{y} | \phi) p(\phi) d\phi} \frac{1}{\int_S p(\phi) d\phi}. \quad (7)$$

Here ϕ is a vector containing all the parameters of (1), $p(\cdot)$ denotes either the posterior or prior density function with support Φ , and S is the set of values of ϕ such that the constraints defined in (5) are satisfied. The first term on the right hand side of (7) equals (6), whereas the second term can be easily computed from draws from the prior.

To assess the plausibility of the given inequality constraints, we calculate the probabilities of each of the columns of Θ_0 satisfying the inequality constraints, and then use (7) to compute the corresponding Bayes factors. If only one of the Bayes factors is greater than 3.2, there is substantial evidence in favour of the corresponding shock being the shock of interest. However, multiple Bayes factors may exceed the threshold, in which case, we proceed by comparing the corresponding constrained models.⁵ For example, if two Bayes factors are greater than 3.2, i.e., two shocks satisfy the constraints with high probability, we calculate the Bayes factor comparing the corresponding constrained models, and if it indicates substantial evidence in favour of one of the constrained models against the other, we label the corresponding shock accordingly.⁶ Otherwise, we conclude that additional

⁵The analogous problem of multiple shocks potentially satisfying the same restrictions is common in the sign restriction literature. For instance, Uhlig (2005) was concerned about a money demand shock potentially satisfying the same sign restrictions as the monetary policy shock, and our empirical results related to this model in section Single monetary policy shock suggest that there indeed are two shocks that satisfy the same inequality constraints, suggesting that the conventional approach to sign restrictions might be unable to identify the monetary policy shock of interest.

⁶To gauge the ability of the Bayes factor to pick the correct shock in finite samples, we conducted a small Monte Carlo simulation experiment. In particular, we checked how efficient it is in labelling the cost (supply) shock in the simple market model, i.e. the SVAR(1) model of Fry and Pagan (2011) (with t -distributed errors) discussed in section II. To this end, we first computed the posterior probability of each of the statistically identified shocks satisfying the inequality constraints that define the supply shock (and the other shock not satisfying them), and then calculated the Bayes factor comparing the corresponding constrained models. We then counted the share of the Monte Carlo

information is needed to discriminate between the plausible shocks. It may come in the form of quantitative information about the likely magnitude of the impulse responses (see, for example, Kilian and Murphy, 2012), as illustrated in section Single monetary policy shock.

It is, of course, possible that none of the shocks satisfies the inequality constraints. In particular, if the Bayes factors comparing the constrained and unconstrained models are smaller than 3.2 for all $k \in \{1, \dots, n\}$, we conclude that the data do not contain sufficient information to discriminate between the plausible shocks, and therefore, these constraints are not helpful in identifying the structural shock of interest. In this case, an alternative set of constraints, potentially arising from competing economic theories, could be considered instead.⁷

Following the sign restriction literature, one may also extend inequality constraints to longer lags in the impulse responses beyond the impact effect, to which our framework also lends itself in a straightforward manner. To that end, in order to impose the same constraints on a single shock embodied in matrix R on $\Theta_j, j = 0, 1, \dots, q$, we redefine the set Q as

$$Q = \{\theta_k : (I_{q+1} \otimes R) \theta_k \geq 0_{J(q+1) \times 1}\}, \quad (8)$$

where θ_k denotes the k th column (corresponding to shock ε_{kt}) of $\Theta = [\Theta'_0, \dots, \Theta'_q]'$, a matrix consisting of the first $q + 1$ structural impulse responses. Analogously to (6), the conditional probability of ε_{kt} being the structural shock of interest is then defined as

$$\Pr(\theta_k \in Q, \theta_{m \neq k} \in Q^c | \mathbf{y}), \quad (9)$$

and the analysis proceeds as in the case of restrictions on Θ_0 only.

Although less common in the sign restriction literature, our framework also accommodates different inequality constraints on different lags. For instance, analogously to Inoue and Kilian (2013), we could impose additional constraints on the sixth lag in Uhlig (2005) SVAR model, when identifying the monetary policy shock. In that case, we would simply replace the matrix $I_{q+1} \otimes R$ in (8) by a block-diagonal matrix $\text{diag}(R^0, \dots, R^q)$, where R^j incorporates the constraints on the j th lag. If not all impulse responses with lags up to q , but only lags belonging to some set L are constrained, then this block diagonal matrix has only the R^j matrices with $j \in L$ on its main diagonal, and the matrix Θ is adjusted accordingly.

The proposed procedure can be summarized as follows:

Algorithm 1.

- Step 1. Estimate the joint posterior distribution of the parameters of the unrestricted SVAR model (1), and compute the posterior distribution of the reduced-form

draws, where this Bayes factor was greater than 3.2. According to the simulation results, with a number of different (reasonable) values of B , and the degree-of-freedom parameters λ_1 , and λ_2 , our method performs well in samples with only 250 observations, and, as expected, its performance improves with increasing sample size (the detailed results, based on 5,000 replications, are not shown to save space, but they are available upon request). For instance, with $b_{11} = b_{22} = 1.2$, $b_{12} = 0.9$, $b_{21} = -1$, $\lambda_1 = \lambda_2 = 5$, the supply shock was correctly labelled in approximately 86% and 98% of the simulated realizations with 250 and 500 observations, respectively, and in almost all cases (99.9%) with 1,000 observations. In the remaining realizations, the Bayes factor is unable to discriminate between the shocks because neither or both of them are found to satisfy the constraints with high probability.

⁷ Alternatively, if we only want to find out whether there is overall evidence in favour of the inequality constraints, we may compute the Bayes factor based on the probability that one or more columns of Θ_0 satisfy the constraints. This probability can be easily calculated by summing the probabilities in (6) over $k \in \{1, \dots, n\}$, because these probabilities are assigned to disjoint events (i.e., they can occur only separately).

impulse response matrices $\Psi_j, j \in L$. If inequality constraints are imposed on all the $q + 1$ first lags of the impulse response function, $L = \{0, 1, \dots, q\}$.

- Step 2. Given the posterior output of B from Step 1, rearrange the columns of each B to ensure that all the posterior draws of B represent the same permutation. This is accomplished by first computing the transformed matrices \tilde{B} , whose each column has Euclidean norm one, and then finding a permutation matrix P for which $C = \tilde{B}P = (c_{ij})$, satisfies $|c_{ii}| > |c_{ij}|$ for all $i < j$. Then, for each B and Ψ_j , the uniquely identified structural impulse responses are given by $\Theta_j = \Psi_j B P D, j \in L$, where D is a diagonal matrix with elements equal to either 1 or -1 that transforms the elements of one of the rows of $B P$ either positive or negative.⁸ As discussed above, normalizing a row corresponding to a variable whose impact effect is restricted by the inequality constraints, facilitates checking the restrictions (expressed for positive shocks) irrespective of the signs of the shocks.
- Step 3. Calculate the probabilities in (9) (or (6)) for all $k \in \{1, \dots, n\}$ using the posterior distribution of the identified structural impulse responses,⁹ and compute the corresponding Bayes factors. If the Bayes factors comparing the constrained and unconstrained models are smaller than 3.2, or, in other words, if the posterior probability of the SVAR model satisfying none of the inequality constraints is high, conclude that the data are not compatible with the constraints, and they cannot thus be used for identification. If only one of the Bayes factors exceeds 3.2, label the corresponding shock as the shock of interest. Otherwise, compare the shocks supported by the data using Bayes factors between them. If only one of these Bayes factors is greater than 3.2, label the corresponding shock the shock of interest. Otherwise conclude that additional information is needed to discriminate between the shocks.

Multiple structural shocks

The procedure introduced in section Single structural shock, generalizes in a straightforward manner to the case of $g > 1$ structural shocks, each of which is restricted by $J_i, i = 1, \dots, g$, inequality constraints. Instead of a single R matrix, we then have g $J_i \times n$ matrices R_i , each embodying the J_i constraints, and the set

$$Q_i = \{ \theta_k : (I_{q+1} \otimes R_i) \theta_k \geq 0_{J_i(q+1) \times 1} \}, \quad (10)$$

contains the columns of the matrix of impulse responses Θ that satisfy the i th inequality constraints.¹⁰

⁸ This sequence of transformations constrains the permissible values of the matrix B to the set \mathcal{B} , defined such that if $B, B^* \in \mathcal{B}$, then necessarily $B = B^*$. The fact that this scheme uniquely identifies the shocks, is shown in appendix A of Lanne *et al.* (2017).

⁹ In practice, this entails computing, for each shock, the share of all draws that satisfy the inequality constraints. The sign patterns are defined for positive shocks, but because the matrix D in Step 2 was defined such that it transforms one of the rows of the permuted impact matrix positive (negative), the share of all draws that satisfy the inequality constraints would be the same for negative shocks.

¹⁰ For notational convenience, we concentrate on the case of the same inequality constraints on each of the g shocks at lags from 0 to q . However, as in the case of a single structural shock, the approach generalizes in a straightforward manner to the case where impulse responses of all shocks are not constrained at all lags, or the constraints on (some of) the shocks are different across the lags.

Analogously to the case of a single shock, computing the posterior probability of the g shocks identified by the inequality constraints calls for going through all combinations of the columns of Θ . For example, the posterior probability of the constrained SVAR model in which the inequality constraints concern two structural shocks ($g = 2$) is given by

$$\Pr(\theta_k \in Q_1, \theta_l \in Q_2, \theta_{m \neq k, l} \in Q_2^c | \mathbf{y}) \text{ for } k, l \in \{1, \dots, n\}, k \neq l, \quad (11)$$

where Q_2^c is the complement of the union $Q_1 \cup Q_2$. In this case, we have $n(n-1)$ different SVAR models to go through. For fixed k and l , (11) is the posterior probability of ε_{kt} and ε_{lt} being the two structural shocks, and the Bayes factor comparing the corresponding constrained and unconstrained models can be calculated from this probability using (7). Furthermore, the Bayes factor based on the sum of these probabilities over all combinations of k and l can be used to determine whether the data lend support to the inequality constraints, as explained in footnote 7.

In general, we have $n!$ permutations of the columns of Θ , on which the g constraints can be placed. However, once the positions of the g shocks have been fixed, the ordering of the remaining unrestricted columns is irrelevant for the assessment of the plausibility of the constraints. Because there are $(n-g)!$ permutations of these columns, the total number of constrained SVAR models that contain the g shocks in fixed positions is $n!/(n-g)!$. This suggests that the posterior probabilities of the constrained SVAR models, such as those in (11), can be evaluated by first calculating the probabilities of the $n!$ SVAR models where the g constraints are imposed on any g columns of Θ , and then marginalizing over each set of the $(n-g)!$ models where they are imposed on same g columns of Θ to obtain the probabilities of the $n!/(n-g)!$ models.

Formally, all $n!$ possible permutations of the columns of Θ can be obtained as the products ΘP^s for $s \in \{1, \dots, n!\}$, where P^s is an $n \times n$ permutation matrix. The probability that the first g columns of ΘP^s satisfy the g inequality constraints can be expressed as

$$\Pr(\theta_1^s \in Q_1, \dots, \theta_g^s \in Q_g, \theta_{m \in \{g+1, \dots, n\}}^s \in Q_g^c | \mathbf{y}), \text{ for } s \in \{1, \dots, n!\} \quad (12)$$

where Q_g^c is the complement of the union $Q_1 \cup \dots \cup Q_g$. Notice that when all n shocks are identified, (12) reduces to $\Pr(\theta_1^s \in Q_1, \dots, \theta_g^s \in Q_g | \mathbf{y}), \text{ for } s \in \{1, \dots, n!\}$. It can be readily checked that the quantities in (12) are the posterior probabilities of all the $n!$ constrained SVAR models. As pointed out above, the probabilities of each of the constrained $n!/(n-g)!$ SVAR models of interest are then obtained by summing the probabilities of the $(n-g)!$ models in which the g inequality constraints are imposed on the same g columns of Θ .

Each of these $n!/(n-g)!$ models represents one ordering of the g structural shocks in the vector ε_t . To find out to which orderings the data lend support, we first compute the Bayes factors of the constrained models corresponding to each one of them against the unconstrained model. If only one of these Bayes factors exceeds 3.2, we have substantial evidence in favour of the corresponding ordering, and can thus label the shocks accordingly. However, multiple Bayes factors may be greater than 3.2, and then we compare the corresponding models by their Bayes factors against each other (in the same way as in the case of one structural shock in section Single structural shock).

If all the Bayes factors are smaller than 3.2, there is little evidence in favour of the inequality constraints as a whole. However, some of them might still be useful, i.e. we might still be able to label some of the structural shocks. To single out the constraints

supported by the data, we recommend a procedure, where some of the shocks are recursively dropped. The probabilities of each set Q_i , ($i = 1, \dots, g$) and the corresponding Bayes factors against the unconstrained model are first computed.¹¹ If all of these Bayes factors are less than 3.2, none of the constraints is useful. Otherwise, the constraints corresponding to the Bayes factors less than 3.2 are relaxed, and then we proceed by comparing the remaining constraints in the same way as in the case of the full set of inequality constraints.¹²

Our recommended procedure discussed above and illustrated by means of an empirical application in section Multiple economic shocks is summarized as Algorithm 2 below.

Algorithm 2.

- Step 1. Follow Steps 1 and 2 of Algorithm 1, to obtain the identified structural impulse responses.
- Step 2. Based on the posterior distribution of the identified structural impulse responses Θ , calculate the probabilities given in (12), and the corresponding Bayes factors using (7). Provided the Bayes factors are not negligible (i.e., at least one of them is greater than 3.2), use them to find the likeliest model (in the same way as in the case of one structural shock in section Single structural shock). Only if all these Bayes factors are smaller than 3.2, go to step 3 below.
- Step 3. Calculate the sum of the probabilities given in (12) for all Q_1, \dots, Q_g individually, and compute the corresponding Bayes factors using (7). If the Bayes factor of the i th constrained model against the unconstrained model is smaller than 3.2, relax the i th constraint, i.e. remove Q_i from Q_1, \dots, Q_g . If all Bayes factors are smaller than 3.2, conclude that the data are not compatible with the inequality constraints. In contrast, if all the Bayes factors are greater than 3.2, remove Q_i corresponding to the smallest Bayes factor. In order to label the remaining shocks, calculate the probabilities given in (12) based on the remaining sets Q_j ($j \neq i$), and compute the corresponding Bayes factors using (7). Provided the Bayes factors are not negligible (i.e., at least one of them is greater than 3.2), use them to find the likeliest model in the same way as with the full set of inequality constraints. If all the Bayes factors are smaller than 3.2, proceed by removing the shock with the greatest Bayes factor against the unconstrained model among the remaining shocks.

IV. Empirical illustrations

We illustrate the methods by means of two empirical applications. The first one, discussed in section Single monetary policy shock is concerned with the computationally most straightforward case of only one shock identified by inequality constraints. In particular, we focus on the monetary policy shock in Uhlig's (2005) model. Our second application in section Multiple economic shocks, in turn, involves multiple sign-identified shocks in Kilian's (2009) model of the crude oil market.

¹¹The probability that one or more shocks satisfy the inequality constraints defined by Q_i can be calculated by summing the probabilities in (12), and the corresponding Bayes factor can be computed from this probability using (7) (see footnote 7).

¹²If all Bayes factors corresponding to individual shocks are greater than 3.2, we recommend dropping the one with the smallest value, for which the evidence is weakest.

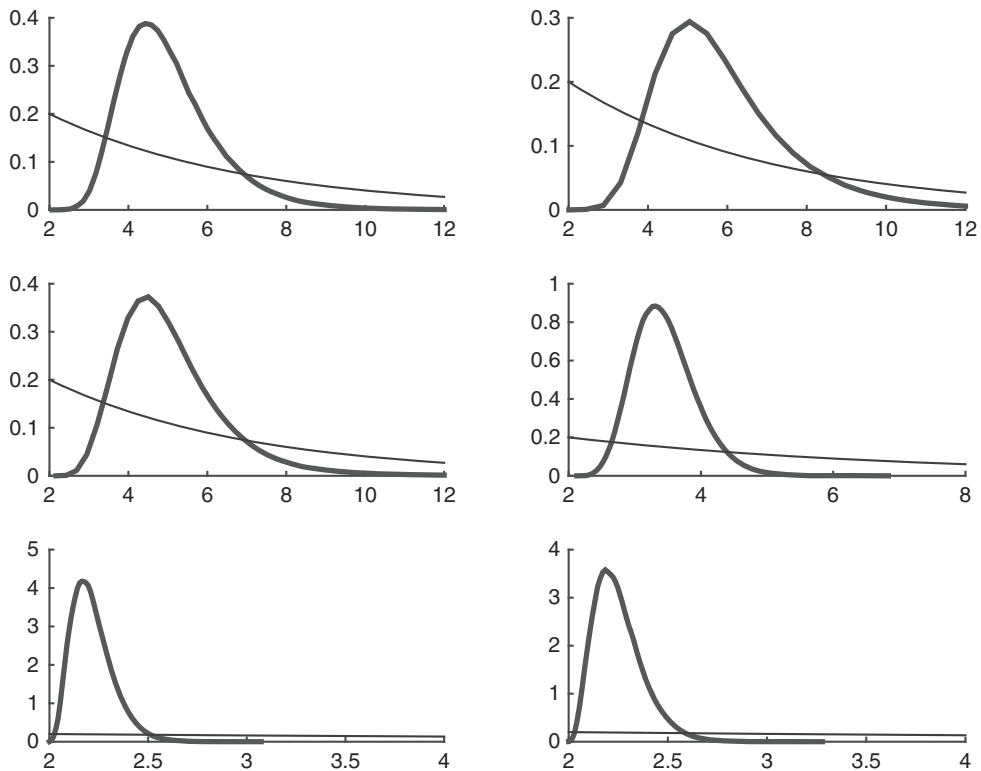


Figure 1. Prior (thin line) and posterior densities of λ_i in Uhlig's (2005) model.

In both applications, we assume that the i th independent component of the error vector of the structural VAR model (1) follows a univariate Student's t distribution with zero mean, unit variance, and λ_i degrees of freedom. This deviates from the Bayesian SVAR literature, where the error vector is typically assumed multivariate normal with a diagonal covariance matrix. It is important to realize that our distributional assumption encompasses the Gaussian case because a t -distributed random variable approaches Gaussianity as the number of degrees of freedom goes to infinity. It is also because of this property of the t distribution that the estimates of λ_i , $i = 1, \dots, n$, indicate the strength of identification (recall that matrix B in model (1) is uniquely identified (up to permutation and scaling of its columns) only under non-Gaussianity of at least $n - 1$ components).

The plots of the prior and posterior densities of λ_i in Uhlig's (2005); Kilian's (2009) models depicted in Figures 1 and 2, respectively, indicate that the error distributions are indeed fat-tailed. In particular, in all cases, the posterior distributions are centred around relatively small values of λ_i , and the data information can be seen to dominate the prior information. These results are based on an exponential prior (truncated such that $\lambda_i > 2$) with mean 7 and variance 25 for each degree-of-freedom parameter λ_i .

As to the error impact matrix B , we operate on its inverse $\text{vec}(B^{-1}) \equiv b$, and assume a Gaussian prior distribution, i.e., $b \sim N(\mathbf{b}, \underline{V}_b)$. The reported results are based on the special case of $\mathbf{b} = \mathbf{0}_{n^2}$, and $\underline{V}_b = c_b I_{n^2}$ with $c_b = 1000^2$, which results in an uninformative, but nevertheless proper prior for B^{-1} . However, the results remain intact whether the

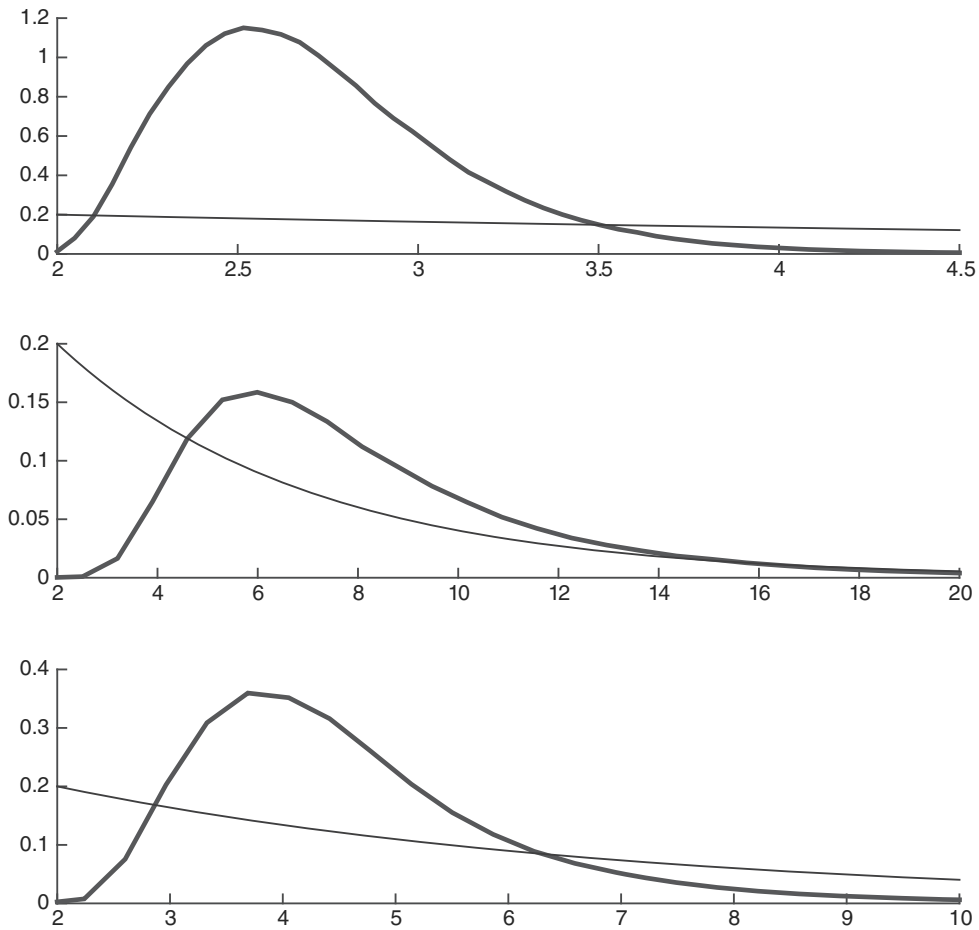


Figure 2. Prior (thin line) and posterior densities of λ_i in Kilian's (2009) model.

(reasonably) informative or an uninformative prior is used (we checked them using $c_b = 10^2$, $c_b = 100^2$, $c_b = 10000^2$, and $c_b^{-1} = 0$ (improper prior)).¹³

We collect the deterministic terms and coefficient matrices of model (1) in matrix $A = [a, A'_1, \dots, A'_p]'$, and assume a multivariate normal prior for $\text{vec}(\mathbf{A}) \equiv \mathbf{a}$, $\mathbf{a} \sim N(\mathbf{a}, \underline{V}_a)$. Following the Bayesian VAR literature, we assume a diagonal prior covariance matrix \underline{V}_a , and set the standard deviation of the (i, j) element of the l th ($l = 1, \dots, p$) coefficient matrix A_l at κ_1/l^{κ_3} for $i = j$, and at $\sigma_i \kappa_1 \kappa_2 / \sigma_i l^{\kappa_3}$ for $i \neq j$. Here σ_i is the residual standard error of a p -lag autoregression for the i th variable in y . As for the deterministic terms, their prior standard deviations are given by $\sigma_i \kappa_4$ (in Kilian's (2009) model κ_4 is set at 10 000, whereas in Uhlig's (2005) model κ_4 is implicitly set at zero). We entertained a number of informative

¹³ Notice that because the SVAR model is point-identified, also an improper prior can be used (i.e., the inference can be based solely on the data). However, in that case, the Bayes factors cannot be evaluated, and, hence, our economic identification procedure based on them is not applicable. This is in contrast to the conventional approach in the sign restriction literature, where the models are only set-identified, and the posterior of the structural parameters within the identified set is proportional to the prior (see Baumeister and Hamilton, 2015). As a result, only under an informative prior does there exist a well-defined posterior for B .

priors for $\text{vec}(\mathbf{A})$, where κ_1 varies from 3 to 10000, κ_2 from 0.2 to 1, and $\kappa_3 = 1$, and found the results intact irrespective of the priors used (the reported results are based on $\kappa_1 = 10$ and $\kappa_2 = 1$). In Kilian's model, we set $\mathbf{a} = 0$ (the (log) variables are in first differences). In Uhlig's model, in turn, we set \mathbf{a} such that the prior mean of the coefficient matrix on the first lag, A_1 , is an identity matrix, and the prior means of the other elements of A are all zero. We defer a more detailed discussion on the technical aspects of estimation to Appendix A.

Single monetary policy shock

Uhlig (2005) studied the effects of the U.S. monetary policy shock in a six-variable structural VAR(12) model with no intercept that we take as given. The monthly time series included in the model are the interpolated real GDP, the interpolated GDP deflator, a commodity price index, total reserves, non-borrowed reserves and the federal funds rate, and, for comparability, the sample period is 1965 : 1–2003 : 12 as in Uhlig (2005). Save the federal funds rate, all variables are expressed in logs.¹⁴

Following Uhlig (2005), we identify only the monetary policy shock. The inequality constraints from his Assumption A.1. state that the first six impulse responses of this shock to prices and non-borrowed reserves are non-positive and to the federal funds rate non-negative (i.e. $q = 5$ in the notation of section III). However, we start with constraints implied by the signs of the impact effects only ($q = 0$), and comment later on the case of constraints on multiple lags. In all cases, the variables are included in vector y_t in the order given above, and the 4×6 matrix R in (5) or (8) equals

$$\begin{pmatrix} 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

As the first step, we compute the probabilities (6) of each of the six columns of B satisfying the inequality constraints, and the corresponding Bayes factors for each constrained model against the unconstrained one. The greatest Bayes factor is 6.3, lending overall support to the constraints on the impact effects. In addition, the data lend substantial support to another model (corresponding to a different shock) with Bayes factor 4.7, while the Bayes factors of the other models (shocks) are close to zero. Thus, we have two plausible candidates for a monetary policy shock. The Bayes factor comparing the former to the latter model equals only 1.3, which suggests that inspection of the impulse responses and the forecast error variance decompositions of these two shocks is needed to find out which one of them is more likely to be the monetary policy shock of interest. As also pointed out by Uhlig (2005), a money demand shock might satisfy the same sign restrictions as the monetary policy shock. If that is indeed the case, the conventional approach to sign restrictions would yield a linear combination of the two shocks, while our approach, by construction, produces two separate shocks.

Based on the impulse responses and the forecast error variance decompositions, we deem the shock associated with 4.7 Bayes factor the monetary policy shock. The impulse

¹⁴ See Uhlig (2005) for a more detailed description of the data set. The data were downloaded from the Estima website at https://estima.com/resources_indx.shtml.

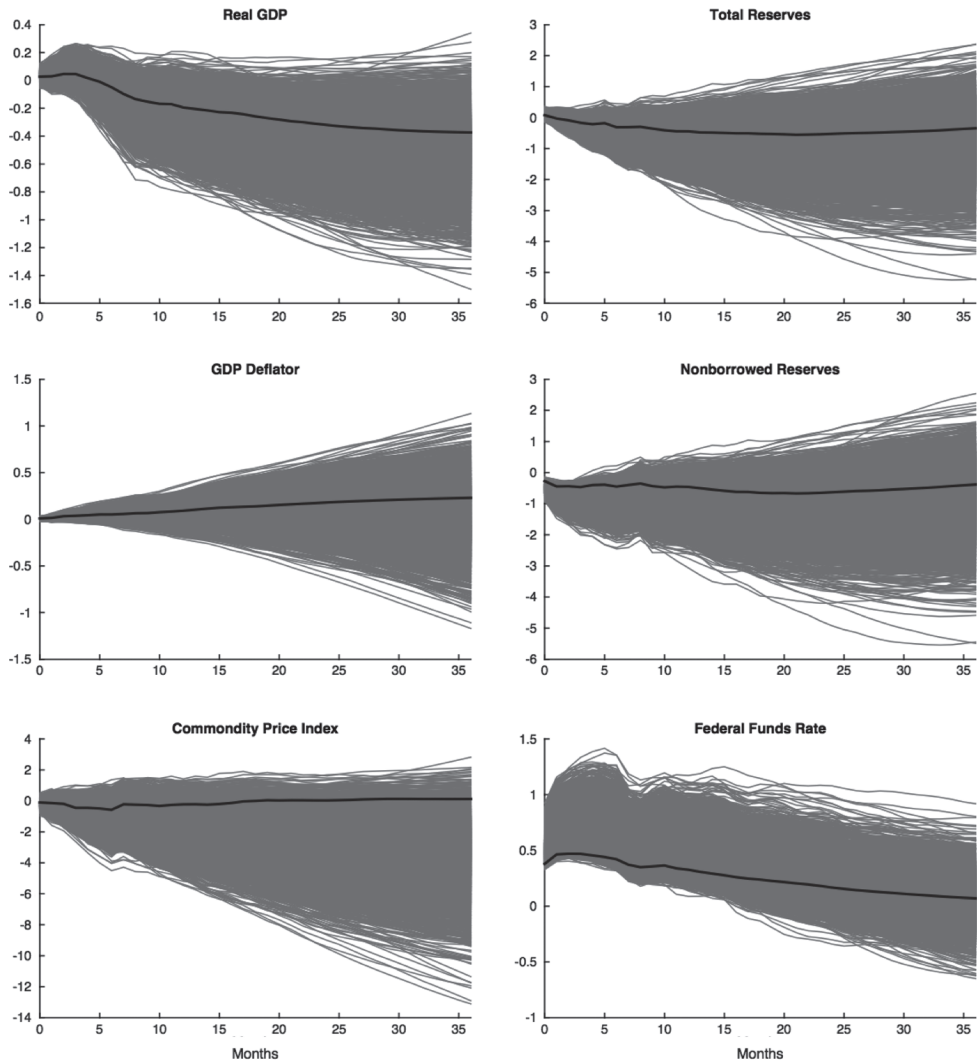


Figure 3. Impulse responses of the contractionary monetary policy shock. The black lines depict the modes of the responses, and the shaded areas are the 68% joint regions of high posterior density.

responses of the contractionary monetary policy shock along with their 68% joint regions of high posterior density (HPD) are depicted in Figure 3. Compared to the results of Inoue and Kilian (2013) based on Uhlig's (2005) original sign restrictions, our impulse responses seem more precisely estimated. As to the response of the real GDP to the monetary policy shock that Uhlig (2005) was mostly interested in, its mode more or less equals zero in the first three months and then turns persistently negative, which is intuitively appealing. While there still remains uncertainty about the effects of monetary policy in that the 68% HPD credible set contains positive output responses, it is the negative values that dominate, in contrast to what Uhlig (2005); Inoue and Kilian (2013) found. This is likely to follow from the fact that our model is exactly identified, whereas sign restrictions alone only reach set

TABLE 1

The relative contributions of the monetary policy shock to the forecast error variances in Uhlig's (2005) model.

Variable	Horizon (months)					
	1	2	6	12	24	36
Real GDP	0.01 (0.00,0.03)	0.01 (0.00,0.04)	0.02 (0.01,0.05)	0.08 (0.03,0.20)	0.24 (0.08,0.47)	0.38 (0.15,0.63)
GDP Deflator	0.00 (0.00,0.01)	0.00 (0.00,0.02)	0.03 (0.00,0.09)	0.03 (0.00,0.12)	0.03 (0.00,0.14)	0.03 (0.00,0.15)
Commodity Price Index	0.00 (0.00,0.02)	0.01 (0.00,0.03)	0.03 (0.00,0.10)	0.07 (0.01,0.20)	0.14 (0.03,0.35)	0.20 (0.04,0.44)
Total Reserves	0.00 (0.00,0.02)	0.00 (0.00,0.01)	0.02 (0.00,0.07)	0.06 (0.01,0.016)	0.09 (0.01,0.25)	0.09 (0.01,0.29)
Non-borrowed Reserves	0.06 (0.03,0.12)	0.08 (0.04,0.16)	0.13 (0.06,0.26)	0.14 (0.06,0.28)	0.14 (0.04,0.32)	0.14 (0.04,0.34)
Federal Funds Rate	0.99 (0.99,1.00)	0.98 (0.97,0.99)	0.93 (0.87,0.96)	0.80 (0.67,0.89)	0.65 (0.47,0.80)	0.57 (0.39,0.76)

The figures are the medians of the posterior distributions of the proportions of the forecast error variance at each horizon accounted for by the monetary policy shock (the 10% and 90% quantiles of the corresponding posterior distributions in parentheses).

identification (cf. the corresponding results of Inoue and Kilian, 2013 based on sign and recursive restrictions).

In Table 1, we report the relative contributions of the monetary policy shock to the forecast error variance of the variables included in the model. As discussed in Appendix B, the forecast error variance decomposition is problematic in the case of sign restrictions that fail to identify a unique model, and it is thus interesting to compare our results to those of Uhlig (2005). In contrast to his results, we find the monetary policy shock to account for only approximately 3% and 7%–20% of the forecast error variance of the GDP deflator and commodity prices, respectively at longer horizons. It also seems to account for a large fraction of the forecast error variance of the real output at longer horizons, with its relative importance increasing considerably with the horizon. While Uhlig found the fraction of the forecast error variance of the federal funds rate accounted for by the monetary policy shock after six months negligible, our results somewhat surprisingly suggest that it is of great importance also at longer horizons although its relative importance diminishes with the horizon as would be expected.

The model (shock) with Bayes factor equal to 6.3 against the unconstrained model may be labelled the money demand shock. Its impulse responses in Figure 4 lie very close to zero at horizons up to 36 months for all variables except the reserves, on which it has a negative effect. Further evidence in favour of this shock being a money demand shock is given by its contributions to the forecast error variances (not shown). It is of relatively minor importance for all variables except the total reserves and non-borrowed reserves, whose forecast error variances it seems to dominate, with contributions ranging between 75.6% and 98.3% for the total reserves and between 28.0% and 64.2% for the non-borrowed reserves, depending on the horizon.

Finally, we also checked Uhlig's (2005) original sign restrictions on the first six impulse responses, and again the shocks labelled the money demand and monetary policy shocks

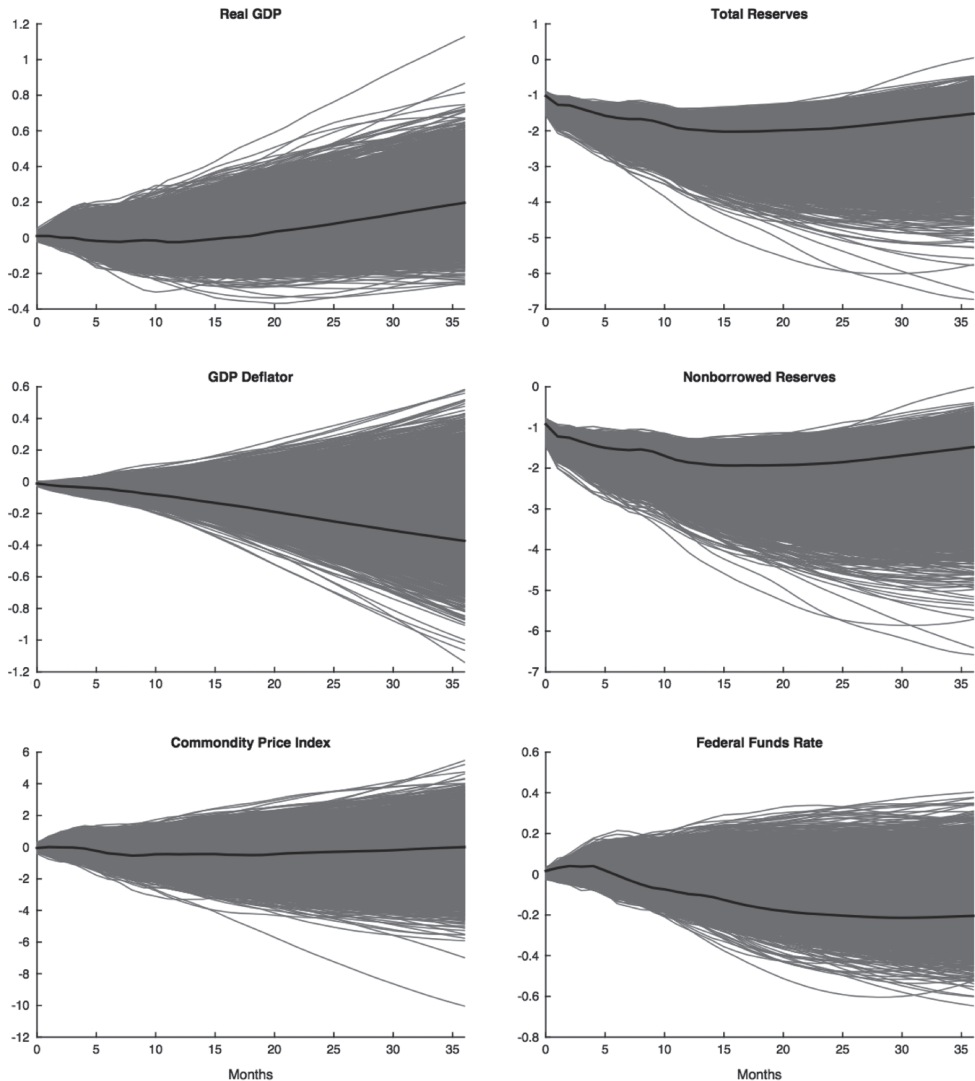


Figure 4. Impulse responses of the money demand shock. The black lines depict the modes of the responses, and the shaded areas are the 68% joint regions of high posterior density.

above turned out the only shocks with Bayes factors greater than 3.2. The Bayes factor comparing the money demand shock to the monetary policy shock equals 7.4, which lends stronger support to the money demand shock. However, in view of the results above, the plausible interpretation also here is that there indeed are two shocks (with different labels) satisfying the same restrictions.

Multiple economic shocks

In this section, we demonstrate our method in the case of multiple shocks of interest. We revisit Kilian's (2009) structural VAR model for the crude oil market that has three variables: percent changes in global oil production (Δoil_t), a business cycle index of global

TABLE 2
Inequality constraints for Kilian's (2009) model.

	<i>Oil production</i>	<i>Real activity</i>	<i>Oil price</i>
Oil supply shock	—	—	+
Aggregate demand shock	+	+	+
Oil-market specific demand shock	+	—	+

The + and — signs indicate that the effect can be non-negative and non-positive, respectively.

real activity (Δg_t), and the real price of crude oil (Δp_t).¹⁵ We follow Kilian (2009) and estimate a VAR(24) model with an intercept.

Kilian and Murphy (2012) identified three shocks in Kilian's (2009) model by restricting the signs of their impact effects as shown in Table 2. For instance, following an oil supply shock, both oil production and real activity decrease, while the oil price increases. Collecting the variables in the vector $y_t = (\Delta oit_t, \Delta g_t, \Delta p_t)'$, the matrices embodying the inequality constraints identifying the oil supply and oil-market specific demand shocks in (10) can be written as

$$R_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } R_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

respectively, while the matrix R_2 corresponding to the aggregate demand shock is a 3×3 identity matrix.

First, we assess the plausibility of each of the six permutations of the shocks satisfying the inequality constraints. Two of the permutations turn out to exhibit non-negligible Bayes factors, with values 12.8 and 3.8, and the Bayes factor comparing the model with strongest support to the other model equals 3.4. Hence, we have managed to successfully identify all three structural shocks, and can proceed with the analysis of their effects.

The modes and the 68% joint regions of high posterior density of the impulse responses of the shocks pertaining to the likeliest permutation are depicted in Figure 5. The oil supply shock is associated with a sharp and permanent drop in oil production, but it has no effect on the real activity and the real price of oil. The aggregate demand shock has a positive impact on the real activity and the real price of oil. The latter jumps up on impact to a permanently higher level, while the real activity gradually increases towards its peak, reaches it in eight months, and starts decreasing thereafter. Kilian (2009); Kilian and Murphy (2012); Inoue and Kilian (2013) also report positive oil price and real activity responses to an aggregate demand shock identified by analogous sign restrictions. Somewhat surprisingly, the oil-market specific demand shock has no a posteriori significant effect on the real price of oil although it has a strong negative effect on the real activity. The latter is in line with standard economic intuition, but not with the results in Kilian (2009), where its effect on real activity was found positive. These data were recently also analysed by Lütkepohl and Netšunajev (2014), who informally labelled the three shocks by making use of sign restrictions in the Markov-switching VAR model of Lanne, Lütkepohl and Maciejowska

¹⁵ For a detailed discussion of the variables, see Kilian (2009). The monthly data for the period 1973:2–2008:9 were downloaded from <http://qed.econ.queensu.ca/jae/2014-v29.3/lutkepohl-netsunajev/>.

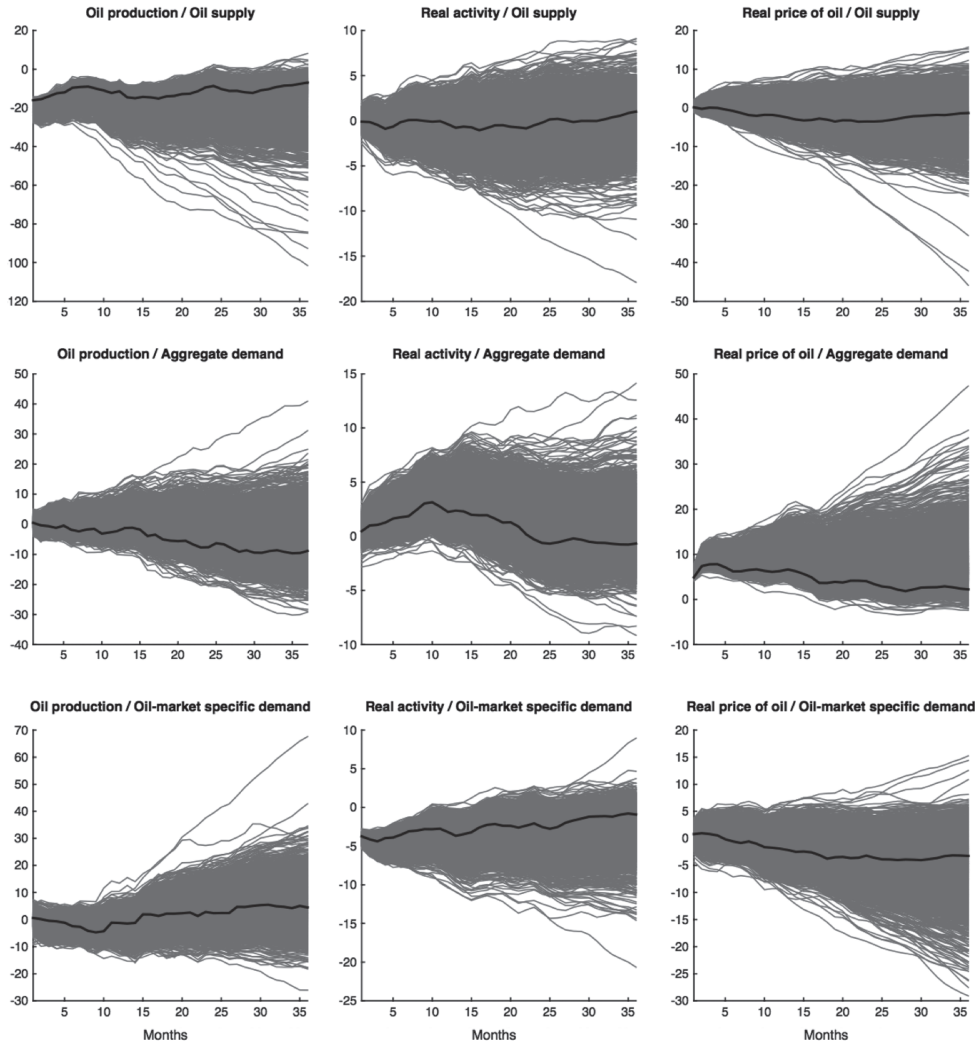


Figure 5. Impulse responses of the three shocks. The black lines depict the modes of the responses, and the shaded areas are the 68% joint regions of high posterior density.

(2010). However, while their results did not clearly object the sign restrictions, the shocks did not seem very strongly identified. Some of their impulse responses were also different from ours. In particular, their results indicated zero impact of the aggregate demand shock on the oil price, whereas our Figure 5 indicates a relatively large positive effect.

In Table 3, we report the relative contributions of the three shocks to the forecast error variance of the real price of oil at selected horizons. The aggregate demand shock seems to account for the bulk of the forecast error variance of oil price at all horizons. This is qualitatively in line with the conclusions of Kilian (2009); Lütkepohl and Netšunajev (2014), who also found the aggregate demand shock important compared to the oil supply shock. However, compared to their results, ours suggest that it has even greater relative importance.

TABLE 3

The relative contributions of the shocks to the forecast error variance of the real price of oil in Kilian's (2009) model.

	<i>Horizon (months)</i>					
	<i>1</i>	<i>2</i>	<i>6</i>	<i>12</i>	<i>24</i>	<i>36</i>
Oil supply shock	0.01 (0.00,0.03)	0.00 (0.00,0.02)	0.01 (0.00,0.06)	0.02 (0.01,0.09)	0.05 (0.01,0.18)	0.06 (0.01,0.23)
Aggregate demand shock	0.98 (0.92,1.00)	0.98 (0.93,1.00)	0.97 (0.90,0.99)	0.94 (0.85,0.98)	0.85 (0.64,0.95)	0.77 (0.50,0.92)
Oil specific demand shock	0.01 (0.00,0.06)	0.01 (0.00,0.06)	0.01 (0.00,0.06)	0.03 (0.01,0.08)	0.07 (0.02,0.23)	0.13 (0.02,0.36)

See notes to Table 1. The figures are the medians of the posterior distributions of the proportions of the forecast error variance at each horizon accounted for by the monetary policy shock (the 10% and 90% quantiles of the corresponding posterior distributions in parentheses).

V. Conclusion

We have introduced a new Bayesian procedure for using inequality constraints implied by economic theory or institutional knowledge to identify economic shocks without imposing any restrictions on the parameters of the structural VAR model. Our procedure is based on the structural VAR model where, following Lanne *et al.* (2017), non-Gaussian and mutually independent errors are assumed. Under these assumptions, the structural shocks, and, hence, their impulse responses are (locally) uniquely identified, which also facilitates checking the validity of any set of sign (inequality) constraints in a straightforward manner. Our contribution is hence twofold. First, we introduce a formal Bayesian procedure to identify economic shocks. The new procedure is less restrictive than the alternative approaches in the macroeconomic SVAR literature in that it does not require any parameter restrictions. Second, we show how the plausibility of inequality constraints can be quantified. Our methods can thus be seen as a formalization of the approaches proposed in the previous statistical identification literature (see, in particular, Lütkepohl and Netšunajev, 2014).

The impulse responses and forecast error variance decompositions of the economic shocks that are found identified with high probability, can then be computed using any of the conventional methods put forth in the literature. Having a uniquely identified SVAR model brings about two great advantages. First, the computations are, in general, much simpler than in the sign identification literature. Second, we avoid the so-called model identification problem arising from the fact that imposing sign restrictions only achieves set identification. This facilitates straightforward interpretation of forecast error variance decompositions and reporting the results of impulse response analysis.

We illustrated the new methods by means of two empirical applications. In Uhlig's (2005) U.S. data set, we found two shocks that satisfy the inequality constraints implied by his sign restrictions for the monetary policy shock. Because our approach, by construction, produces two separate shocks in this case, we were able to distinguish between them, unlike the conventional approach. While there was great uncertainty about the impact of the (contractionary) monetary policy shock on the real GDP, we found its effect negative after the first few quarters, which is intuitively appealing. In Kilian's (2009) model of the

crude oil market, we were able to convincingly identify all three shocks by making use of the inequality constraints implied by the sign restrictions of Kilian and Murphy (2012).

Our procedure could be extended to checking the validity of and discriminating between alternative dynamic stochastic general equilibrium (DSGE) models. Recently, Canova and Paustian (2011) suggested a two-step procedure where a set of robust (sign) restrictions implied by a DSGE model (or multiple competitive DSGE models) is first imposed to identify a SVAR model, and the plausibility of another set of restrictions is then checked in this identified model. Our procedure could be used in both steps for checking the plausibility of the inequality constraints implied by the DSGE models and finding the maximal set of robust constraints. Furthermore, the probabilities of the second-step constraints could subsequently be computed conditional on the shocks identified in the first step.

Appendix A

In this appendix, we describe the Metropolis-within-Gibbs algorithm used for the estimation of the posterior distribution of the parameters of the SVAR model in (1).

Let us start by describing the conditional likelihood function. We assume that the i th component ($i \in \{1, \dots, n\}$) of the error ε_t follows Student's t distribution with λ_i degrees of freedom. For computational convenience, we reparametrize ε_{it} as $h_{it}^{-1/2}\eta_{it}$, where η_{it} is a standard normal random variable, and $(\lambda_i - 2)h_{it}$ follows the chi-square distribution with λ_i degrees of freedom ($\lambda_i > 2$).¹⁶ Then, $\varepsilon_t = H_t^{-1/2}\eta_t$, where η_t is a $(n \times 1)$ vector of standard normal random variables, and $H_t = \text{diag}(h_{1t}, \dots, h_{nt})$. From (1), a change of variable yields

$$p(\mathbf{y} | \mathbf{A}, \mathbf{B}, \mathbf{H}) \propto |\det(\mathbf{B}^{-1})|^T \prod_{t=1}^T |H_t|^{1/2} \exp \left[-\frac{1}{2} \sum_{t=1}^T \mathbf{u}_t' \mathbf{B}^{-1} H_t \mathbf{B}^{-1} \mathbf{u}_t \right], \quad (\text{A.1})$$

where $\mathbf{A}' = [a, A_1, \dots, A_p]$, $\mathbf{H} = \text{diag}(h_{11}, \dots, h_{n1}, \dots, h_{1T}, \dots, h_{nT})$, $\mathbf{u}_t = \mathbf{y}_t - \mathbf{a} - A_1 \mathbf{y}_{t-1} - \dots - A_p \mathbf{y}_{t-p}$, and \mathbf{y} is a $(Tn \times 1)$ vector obtained by stacking \mathbf{y}_t for $t = 1, \dots, T$.

We operate on \mathbf{B}^{-1} , the inverse of \mathbf{B} , and to facilitate unique identification, we make two additional assumptions. First, we restrict the parameter space of $\mathbf{B}^{-1} = (c_{ij})$ such that $|c_{ij}| > |c_{ji}|$ for all $i > j$. Second, we assume that the diagonal elements of \mathbf{B}^{-1} are positive. In practice, these conditions are imposed by multiplying the conditional likelihood by an indicator function, which equals unity if \mathbf{B}^{-1} belongs to the defined space, and zero otherwise. Notice that because the likelihood function (and therefore the posterior) is invariant with respect to permutation of the rows of \mathbf{B}^{-1} (columns of \mathbf{B}), we can reorder the columns of the restricted posterior \mathbf{B} matrices produced by Markov chain Monte Carlo simulation without changing the posterior model probabilities (see Geweke, 2007 for discussion).

A Gaussian prior distribution is assumed for $\text{vec}(\mathbf{B}^{-1}) = \mathbf{b}$, $\mathbf{b} \sim N(\underline{\mathbf{b}}, \underline{\mathbf{V}}_{\mathbf{b}})$, and we simulate from the conditional posterior of \mathbf{B}^{-1} by an accept-reject Metropolis-Hastings (ARMH) algorithm (see, for example, Chib and Greenberg, 1995). To obtain a good

¹⁶ It can be readily shown that with this reparametrization, the density kernel of the distribution of the product $h_{it}^{-1/2}\eta_{it}$ is equal to that of ε_t under the unit variance.

proposal density for B^{-1} , we approximate the log conditional likelihood by the second order Taylor expansion around some $\tilde{\mathbf{b}}$:

$$\log p(\mathbf{y} | \mathbf{A}, B, \mathbf{H}) \approx \log p(\mathbf{y} | \mathbf{A}, \tilde{\mathbf{b}}, H) \\ + (\mathbf{b} - \tilde{\mathbf{b}})' \mathbf{f} - 0.5 (\mathbf{b} - \tilde{\mathbf{b}})' \mathbf{G} (\mathbf{b} - \tilde{\mathbf{b}}),$$

where \mathbf{f} and \mathbf{G} are the gradient and the negative of the Hessian of the log conditional likelihood evaluated at $\tilde{\mathbf{b}}$, respectively. Combining the above with the prior density yields

$$\log p(\mathbf{b} | \mathbf{A}, \mathbf{H}, \mathbf{y}) \approx -\frac{1}{2} [\mathbf{b}' (\mathbf{G} + \underline{V}_{\mathbf{b}}^{-1}) \mathbf{b} - 2\mathbf{b}' (\mathbf{f} + \mathbf{G}\tilde{\mathbf{b}} + \underline{V}_{\mathbf{b}}^{-1}\tilde{\mathbf{b}})],$$

which is a (log) kernel of a multivariate normal density. We construct the Taylor expansion around the mode, $\tilde{\mathbf{b}} = \hat{\mathbf{b}}$. At first, we make no additional assumptions concerning the space of B^{-1} , and thus the (local) posterior mode can be quickly obtained by the Newton-Raphson method, using explicit formulae for \mathbf{f} and \mathbf{G} (not shown to save space, but available upon request), and the current draw of \mathbf{b} as an initial point (see, for example, Chan, 2015).

If the resulting local mode does not satisfy the restrictions stated below the likelihood function (A.1), we replace $\hat{\mathbf{b}}$ and \mathbf{G} (evaluated at $\hat{\mathbf{b}}$) with $\bar{\mathbf{b}} = (I_n \otimes DP)\hat{\mathbf{b}}$ and $\bar{\mathbf{G}} = (I_n \otimes (DP)^{-1})\mathbf{G}(I_n \otimes (DP)^{-1})$, respectively, where P is the permutation matrix for which the matrix $P\widehat{B}^{-1} = (\hat{c}_{ij})$ satisfies $|\hat{c}_{ij}| > |\hat{c}_{ji}|$ for all $i > j$, and D is a diagonal matrix with elements equal to either 1 or -1 that transforms the diagonal elements of $P\widehat{B}^{-1}$ positive.

Because the latter transformation may change the signs of the rows of \widehat{B}^{-1} , it may result in a value of the proposal density which is virtually zero at the current draw of \mathbf{b} (causing the proposal, say \mathbf{b}^* , to be rejected). Therefore, to improve the performance of the sampler, we use the following mixture of two multivariate normal densities as a proposal density in the ARMH algorithm:

$$q(\mathbf{b}) = \frac{1}{2} \mathbf{N}(\mathbf{b} | \hat{\mathbf{b}}, \mathbf{G}) + \frac{1}{2} \mathbf{N}(\mathbf{b} | \bar{\mathbf{b}}, \bar{\mathbf{G}}).$$

In our practical implementations of the algorithm, only a few draws from $q(\mathbf{b})$ are typically required in the accept-reject (AR) steps. Furthermore, the Metropolis-Hastings (MH) acceptance rates tend to vary between 0.85% and 0.99%.

As far as the full conditional posterior of \mathbf{A} is concerned, it is easy to check that the conditional likelihood (A.1) can also be expressed as

$$p(\mathbf{y} | \mathbf{A}, B, \mathbf{H}) \propto \exp \left[-\frac{1}{2} (\mathbf{y} - \mathbf{X} \text{vec}(\mathbf{A}))' \boldsymbol{\Omega} (\mathbf{y} - \mathbf{X} \text{vec}(\mathbf{A})) \right],$$

where \mathbf{X} is obtained by stacking $(I_n \otimes X_t)$ for $t = 1, \dots, T$, $X_t = (1, y'_{t-1}, \dots, y'_{t-p})$, and $\boldsymbol{\Omega} = (I_T \otimes B^{-1})H(I_T \otimes B^{-1})$. Assuming a multivariate normal prior for $\text{vec}(\mathbf{A}) = \mathbf{a}$, $\mathbf{a} \sim N(\underline{\mathbf{a}}, \underline{V}_{\mathbf{a}})$, we then obtain

$$\text{vec}(\mathbf{A}) | B, \mathbf{H}, \mathbf{y} \sim \mathbf{N}(\bar{\mathbf{a}}, \bar{V}_{\mathbf{a}}),$$

where $\bar{V}_{\mathbf{a}}^{-1} = \underline{V}_{\mathbf{a}}^{-1} + \mathbf{X}'\boldsymbol{\Omega}\mathbf{X}$, and $\bar{\mathbf{a}} = \bar{V}_{\mathbf{a}}(\underline{V}_{\mathbf{a}}^{-1}\underline{\mathbf{a}} + \mathbf{X}'\boldsymbol{\Omega}\mathbf{y})$. The precision-based sampling method of Chan and Jeliazkov (2009) can be used to simulate draws from $N(\bar{\mathbf{a}}, \bar{V}_{\mathbf{a}})$ efficiently.

We now turn to the sampling of the latent variables $\{h_{1t}, \dots, h_{nt}\}_{t=1}^T$. The log conditional likelihood is proportional to

$$\begin{aligned} \log p(\mathbf{y} | \mathbf{A}, B, \mathbf{H}) &\propto \sum_{t=1}^T \log |H_t|^{1/2} - \frac{1}{2} \sum_{t=1}^T \mathbf{u}_t' B^{-1} H_t B^{-1} \mathbf{u}_t \\ &= \sum_{t=1}^T \left[\sum_{i=1}^n \log h_{it}^{1/2} - \frac{1}{2} \varepsilon_t' H_t \varepsilon_t \right] \\ &= \sum_{t=1}^T \left[\sum_{i=1}^n \left(\log h_{it}^{1/2} - \frac{1}{2} h_{it} \varepsilon_{it}^2 \right) \right], \end{aligned}$$

where $\varepsilon_t = B^{-1} \mathbf{u}_t$ and $\mathbf{u}_t = \mathbf{y}_t - \mathbf{a} - A_1 \mathbf{y}_{t-1} - \dots - A_p \mathbf{y}_{t-p}$. Recall that the (hierarchical) prior of each $(\lambda_i - 2)h_{it}$ is the chi-squared distribution with λ_i degrees of freedom. Then, by multiplying $p(\mathbf{y} | \mathbf{A}, B, \mathbf{H})$ by the product of the prior densities of h_{it} for $t \in \{1, \dots, T\}$ and $i \in \{1, \dots, n\}$ we obtain

$$p(h_{it} | \mathbf{A}, B, \lambda, \mathbf{y}) \propto h_{it}^{(\lambda_i-1)/2} \exp \left\{ - \left[\lambda_i - 2 + \varepsilon_{it}^2 \right] h_{it} / 2 \right\},$$

where $\lambda = (\lambda_1, \dots, \lambda_n)$. This implies that each h_{it} ($t \in \{1, \dots, T\}, i \in \{1, \dots, n\}$) can be sampled from the chi-square distribution as follows:

$$\left[\lambda_i - 2 + \varepsilon_{it}^2 \right] h_{it} | \mathbf{A}, B, \lambda, \mathbf{y} \sim \chi^2(\lambda_i + 1).$$

We assume an exponential prior distribution for each λ_i , $\lambda_i \sim \text{Exp}(\lambda_i)$. From the hierarchical prior density of h_{it} ($t \in \{1, \dots, T\}$) and the assumption $\lambda_i \sim \text{Exp}(\lambda_i)$, it follows that the conditional posterior density of λ_i can be written as proportional to

$$\begin{aligned} p(\lambda_i | \{h_{it}\}_{t=1}^T, \mathbf{y}) &\propto \left[2^{\lambda_i/2} \Gamma(\lambda_i/2) \right]^{-T} (\lambda_i - 2)^{\lambda_i T/2} \left(\prod_{t=1}^T h_{it}^{(\lambda_i-2)/2} \right) \\ &\quad \times \exp \left[- \left(\frac{1}{\lambda_i} + \frac{\lambda_i - 2}{2\lambda_i} \sum_{t=1}^T h_{it} \right) \lambda_i \right]. \end{aligned}$$

It is the hierarchical prior structure in which each λ_i affects the data only through $\{h_{it}\}_{t=1}^T$ that lies behind this result. Following Geweke (2005), we simulate from the conditional posterior of the degree-of-freedom parameter λ_i using an independence-chain MH algorithm. As a candidate distribution of λ_i , we use the univariate normal distribution with mean equal to the mode of the log conditional posterior, and the precision parameter equal to the negative of the second derivative of the log posterior density evaluated at the mode.

Appendix B

In this appendix, we discuss the computation of impulse responses and forecast error variance decomposition of the identified shocks in SVAR model (1). As our model produces unique impulse response functions, conventional pointwise posterior median impulse responses and error bands could be entertained in a straightforward manner. It is however, well known that, while frequently applied, these may also yield misleading conclusions.

Therefore, we recommend employing an extension of the approach of Inoue and Kilian (2013), who derived the joint posterior density of the impulse responses and recommended reporting their mode and $100(1 - \alpha)\%$ highest posterior density (HPD) credible set.

The posterior density of the structural impulse responses implied by our model can be derived in a straightforward manner. For notational simplicity, let us ignore deterministic terms, and collect the coefficient matrices of model (1) in matrix $\tilde{\mathbf{A}} = [A_1 \dots A_p]$. Because the model is exactly identified (statistically), there is a one-to-one mapping between the first $p + 1$ structural impulse responses $\tilde{\Theta} = [B', (\Psi_1 B)', \dots, (\Psi_p B)']'$ and $[B, \tilde{\mathbf{A}}]$, and the nonlinear function $\tilde{\Theta} = f(B, \tilde{\mathbf{A}})$ is known. By a change of variable, the posterior density of the first $p + 1$ structural impulse responses $\tilde{\Theta} = [B', (\Psi_1 B)', \dots, (\Psi_p B)']'$ can thus be written as

$$\begin{aligned} p(\tilde{\Theta}|\mathbf{y}) &= \left| \frac{\partial [\text{vec}(B)', \text{vec}(\tilde{\mathbf{A}})']}{\partial \text{vec}(\tilde{\Theta})} \right| p(B, \tilde{\mathbf{A}}|\mathbf{y}) \\ &= \left(\left| \frac{\partial \text{vec}(\tilde{\Theta})}{\partial [\text{vec}(B)', \text{vec}(\tilde{\mathbf{A}})']} \right| \right)^{-1} p(B, \tilde{\mathbf{A}}|\mathbf{y}) \\ &= |B|^{-np} p(B, \tilde{\mathbf{A}}|\mathbf{y}). \end{aligned} \quad (\text{B.1})$$

where $p(B, \tilde{\mathbf{A}}|\mathbf{y})$ is the joint posterior density of B and $\tilde{\mathbf{A}}$, and the second equality follows by the inverse function theorem. In the terminology of Inoue and Kilian (2013), the model corresponding to a draw $(B, \tilde{\mathbf{A}})$ from the posterior distribution of the parameters of the SVAR model that maximizes (B.1) is the modal model that produces the mode of the structural impulse responses. Note that the above posterior density is defined only for the first $p + 1$ impulse responses, for which the Jacobian can be evaluated analytically.

In addition to the mode, it is useful to have a measure of the uncertainty surrounding the impulse responses, and, following Inoue and Kilian (2013), we define the $100(1 - \alpha)\%$ HPD credible set of the first $p + 1$ impulse responses as

$$S = \left\{ \tilde{\Theta} : p(\tilde{\Theta}|\mathbf{y}) \geq c_\alpha \right\},$$

where c_α is the largest constant such that $\Pr(S) \geq 1 - \alpha$. We then report the impulse responses up to some prespecified horizon of models belonging to this set, in addition to those of the modes. In the empirical literature it is customary to set α equal to 0.32, i.e., to report the 68% credible sets. As Inoue and Kilian (2013) pointed out, there is no reason for these credible sets to be dense, but they will typically exhibit a ‘shot-gun’ pattern.

As far as the forecast error variance decompositions are concerned, they can be calculated in a standard fashion using the mode of the structural impulse responses as defined above (see, for example, Lütkepohl, 2005, Chapter 2.3). For sign-identified models, forecast error variance decompositions based on pointwise median impulse responses are typically reported for sign-identified models. However, as pointed out by Fry and Pagan (2011), they have the problem that they are based on correlated shocks, and may therefore be difficult to interpret because the contributions of all shocks need not sum to unity for all variables. In contrast, since our model is uniquely identified, this problem is avoided. Moreover, in

contrast to the conventional approach to sign restrictions, our approach facilitates analyzing the effects of shocks of a given size, and thus answering questions like ‘what would be the responses to a 25 basis point interest rate shock.’

Final Manuscript Received: July 2019

References

- Antolín-Díaz, J. and Rubio-Ramírez, J. F. (2018). ‘Narrative sign restrictions for SVARs’, *American Economic Review*, Vol. 108, pp. 2802–2829.
- Baumeister, C. and Hamilton, J. D. (2015). ‘Sign restrictions, structural vector autoregressions, and useful prior information’, *Econometrica*, Vol. 83, pp. 1963–1999.
- Canova, F. and De Nicoló, G. (2002). ‘Monetary disturbances matter for business fluctuations in the G-7’, *Journal of Monetary Economics*, Vol. 49, pp. 1131–1159.
- Canova, F. and Paustian, M. (2011). ‘Business cycle measurement with some theory’, *Journal of Monetary Economics*, Vol. 58, pp. 345–361.
- Chan, J. C. C. (2015). ‘The stochastic volatility in mean model with time-varying parameters: an application to inflation modeling’, *Journal of Business and Economic Statistics*, Vol. 35, pp. 17–28.
- Chan, J. and Jeliazkov, I. (2009). ‘Efficient simulation and integrated likelihood estimation in state space models’, *International Journal of Mathematical Modelling and Numerical Optimisation*, Vol. 1, pp. 101–120.
- Chib, S. and Greenberg, E. (1995). ‘Understanding the Metropolis-Hastings algorithm’, *The American Statistician*, Vol. 49, pp. 327–335.
- Faust, J. (1998). ‘The robustness of identified VAR conclusions about money’, *Carnegie-Rochester Conference Series on Public Policy*, Vol. 49, pp. 207–244.
- Fry, R. and Pagan, A. (2011). ‘Sign restrictions in structural vector autoregressions: a critical review’, *Journal of Economic Literature*, Vol. 49, pp. 938–960.
- Geweke, J. (2005). *Contemporary Bayesian Econometrics and Statistics*. Wiley, Hoboken, New Jersey.
- Geweke, J. (2007). ‘Interpretation and inference in mixture models: simple MCMC works’, *Computational Statistics & Data Analysis*, Vol. 51, pp. 3529–3550.
- Giacomini, R. and Kitagawa, T. (2014). *Inference About Non-Identified SVARs*. Cemmap Working Papers 45/14.
- Herwartz, H. and Lütkepohl, H. (2014). ‘Structural vector autoregressions with Markov switching: combining conventional with statistical identification of shocks’, *Journal of Econometrics*, Vol. 183, pp. 104–116.
- Hyvärinen, A., Zhang, K., Shimizu, S. and Hoyer, P. (2010). ‘Estimation of a structural vector autoregressions model using non-Gaussianity’, *Journal of Machine Learning Research*, Vol. 11, pp. 1709–1731.
- Inoue, A. and Kilian, L. (2013). ‘Inference on impulse response functions in structural VAR models’, *Journal of Econometrics*, Vol. 177, pp. 1–13.
- Kass, R. E. and Raftery, A. E. (1995). ‘Bayes factors’, *Journal of the American Statistical Association*, 90, pp. 773–795.
- Kilian, L. (2009). ‘Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market’, *American Economic Review*, Vol. 99, pp. 1053–1069.
- Kilian, L. and Lütkepohl, H. (2017). *Structural Vector Autoregressive Analysis*. Cambridge University Press, Cambridge, UK.
- Kilian, L. and Murphy, D. (2012). ‘Why agnostic sign restrictions are not enough: understanding the dynamics of the oil market VAR models’, *Journal of the European Economic Association*, Vol. 10, pp. 1166–1188.
- Koop, G. (2003). *Bayesian Econometrics*. Wiley, Chichester.
- Lanne, M. and Luoto, J. (2019). ‘GMM estimation of non-Gaussian structural vector autoregression’, *Journal of Business and Economic Statistics*, (forthcoming).
- Lanne, M., Lütkepohl, H. and Maciejowska, K. (2010). ‘Structural vector autoregressions with Markov switching’, *Journal of Economic Dynamics and Control*, Vol. 34, pp. 121–131.
- Lanne, M., Meitz, M. and Saikkonen, P. (2017). ‘Identification and estimation of non-Gaussian structural vector autoregressions’, *Journal of Econometrics*, Vol. 196, pp. 288–304.

- Lütkepohl, H. (1996). *Handbook of Matrices*. Wiley, Chichester.
- Lütkepohl H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin.
- Lütkepohl H. and Netšunajev, A. (2014). ‘Disentangling demand and supply shocks in the crude oil market: how to check sign restrictions in structural VARs’, *Journal of Applied Econometrics* 29, 479–496.
- Piffer, M. (2016). *Assessing Identifying Restrictions in SVAR Models*. DIW Berlin Discussion Papers 1563.
- Rubio-Ramírez, J. F., Waggoner, D. F. and Zha, T. (2010). ‘Structural vector autoregressions: theory of identification and algorithms for inference’, *Review of Economic Studies*, Vol. 77, pp. 665–696.
- Straub, R. and Peersman, G. (2006). *Putting the New Keynesian model to a test*. International Monetary Fund Working Papers 06/135.
- Uhlig, H. (2005). ‘What are the effects of monetary policy on output? Results from an agnostic identification procedure’, *Journal of Monetary Economics*, Vol. 52, pp. 381–419.